UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/514,413	11/15/2004	Nick Campbell	71109-014	6822

20277   7590   05/28/2008
MCDERMOTT WILL & EMERY LLP
600 13TH STREET, N.W.
WASHINGTON, DC 20005-3096

EXAMINER
SHAH, PARAS D

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
05/28/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *29 February 2008*.

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-19* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-19* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All   b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☐ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____ .

5)☐ Notice of Informal Patent Application

6)☐ Other: _____.

DETAILED ACTION

1. This communication is in response to the Amendments and Arguments filed on

02/29/2008. Claims 1-19 are pending and have been examined. The Applicants'

amendment and remarks have been carefully considered, but they are not persuasive

and do not place the claims in condition for allowance. Accordingly, this action has been

made FINAL.

2. All previous objections and rejections directed to the Applicant's disclosure and

claims not discussed in this Office Action have been withdrawn by the Examiner.

Response to Arguments

3.   Applicant's arguments (pages 7-9) filed on 02/29/2008 with regard to claims 1-19

have been fully considered but they are not persuasive.

      In response to the 35 USC 112, 2nd paragraph rejection, the rejection is

maintained as the claims still are unclear as to what the word reliable and un-reliable

means in the context of the limitation. Please see below for the claims and

interpretation.

      In regards to independent claims 1, 8, and 14, the Applicants argue that the

references Lea in view of Mermelstein in view of Schmidbauer fail to teach the

limitations of "a distribution of energy of a prescribed frequency range" and "a

distribution of spectrum" as recited in the first and second paragraphs of claims 1, 8,

and 14. The Examiner respectfully disagrees with respect to all arguments presented. In

regards to the former limitation, the primary reference of Lea does teach "a distribution

of energy of a prescribed frequency range". On page 42.7.1, Figure 1, a speech

waveform is input and energy calculations are made for specific frequency ranges

(prescribed frequency ranges) (sonorant energy filter and very low frequency filter).

Further, on page 42.7.2, left column, sect. 3, 1st full paragraph-right column, a speech

waveform is passed through and energy values are calculated over 30ms windows.

Hence, an energy distribution is done to determine voicing. In regards to the latter

limitation, where "a distribution of spectrum of said speech waveform," is also taught by

Lea. On page, 42.7.1, Figure 1, a speech waveform is input and it is obvious that the

speech is of a specific length. Further, on page 42.7.2, left column sec. 3, 1st full

paragraph-right column, each frame of speech (second portion is the portion after the

first portion (or frame) has been input) is analyzed to determine voicing and also energy

values are being calculated for each window. Hence, a series of windows represent a

spectrum (energy if plotted). Further, with respect to Figure 2 on page 42.7.3, the last

plot, shows an energy spectrum plot over successive frames in order to determine

syllables. Hence, Lea in view of Mermelstein in view of Schmidbauer teach the stated

limitations in the above mentioned claims.

In regards to independent claims 5, 12, 18, the Applicants argue that the

limitation of "a distribution waveform of energy" is not taught by Lea in view of

Mermelstein. The Examiner respectfully disagrees with this argument. On page 42.7.1,

right column, sect. 2, 1st full paragraph, and Figure 1, a prescribed frequency range

energy calculations are made for specific frequency ranges (prescribed frequency

ranges) (sonorant energy filter and very low frequency filter) is used and dips of energy

define minimums where syllables are located based on energy and voicing decision

(see Lea). Further, the secondary reference by Mermelstein teaches the use of

separating and calculating a distribution of energy of a speech using a local minimum of

a time versus loudness plot in order to determine syllables in a speech signal (see

Abstract and page 881, left column, sect. I, Figure 1). The loudness of Mermelstein, is

determined from the speech power by weighting the spectrum in terms of frequency

band (see page 881, left column, sect. I, 1st paragraph. Hence Lea in view of

Mermelstein teach the stated limitations in the above mentioned claims.

Claim Rejections - 35 USC § 112

4. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

5. Claims 1, 7, 8, 10, 13, 14, and 16 are rejected under 35 U.S.C. 112, second

paragraph, as being indefinite for failing to particularly point out and distinctly claim the

subject matter which applicant regards as the invention. Please see below for the

reasons of indefiniteness. Further, the limitations "portion reliably representing" and

"non-reliability" and is not understandable as to what the applicant is seeking to claim.

The mentioned limitations were interpreted to mean ranges where a syllabic nuclei s

extracted and where a voiced region was determined.

6. Claims 2-4, and 9-11 are rejected as being dependent upon an indefinite base claim.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 1, 2, 4, 8, 9, 11, 14, 15, and 17 are rejected under 35 U.S.C. 103(a) as being

unpatentable over in Lea *et al.* ("Algorithms for acoustic prosodic analysis") in view of

Mermelstein ("Automatic segmentation of speech into syllabic units") in view of

Schmidbauer ("Syllable-based Segment-hypotheses Generation in Fluently spoken

speech using Gross Articulatory features.).

    As to claims 1, 8, and 14, Lea *et al.* teaches an apparatus for determining, based

on speech waveform data, a portion reliably representing a feature of the speech

waveform, comprising:

        extracting means for calculating (see Figure 1, sonorant energy filter and

energy calculation), from said data, distribution of an energy of a prescribed

frequency range of said speech waveform on a time axis, and for extracting,

among various syllables, a first portion of said speech waveform (See page

42.7.1, Figure 1, a speech waveform is input and energy calculations are made

for specific frequency ranges (prescribed frequency ranges) (sonorant energy

filter and very low frequency filter)), that is generated stably by a source of said

speech waveform, based on the distribution of energy and pitch of said speech

waveform (see Figure 1) (e.g. From the figure, speech is input into the system.

Then, energy calculation is done to determine the syllable units (voicing).

Further, a stable range is determined from the boundary that is determined by

pitch. (see page 42.7.1, right column, last paragraph-page 42.7.2, left column,

lines 1-12));

estimating means for calculating (See Figure 1, energy calculation), from

said data, distribution of spectrum of said speech waveform on the time axis, and

estimating, based on the distribution of spectrum, a second portion (see page

42.7.2, left column sec. 3, 1st full paragraph-right column, each frame of speech

(second portion is the portion after the first portion (or frame) has been input) is

analyzed to determine voicing and also energy values are being calculated for

each window) of said speech waveform for which change is well controlled by

said source (see Figure 2 and page 42.7.3, right column, 1st full paragraph) (e.g.

In the cited section two types of methods are compared. A speech spectrum is

obtained for both methods in order to determine the boundary for each syllable,

which is well controlled. The well-controlled portions is determined of the

boundary extracted (e.g. reliable));

However, Lea does not specifically teach the minimum of a time

distribution waveform.

Mermelstein does teach the use of a time distribution waveform for

detecting local minimums (see Figure 1, and page 881, left column, sect. I, entire

section) (e.g. The cited section uses a convex-hull to determine local minimum

on loudness versus time waveform.)

It would have been obvious to one of ordinary skilled in the art at the time

the invention as made to have modified the separation of speech signal into

quasi-syllables as taught by Lea with the use of a time-distribution waveform as

taught by Mermelstein. The motivation to have combined the references involves

the segmentation of speech into syllable units (see Abstract).

However, Lea in view of Mermelstein do not specifically teach the range

being stably extracted by the source.

Schmidbauer does teach

means for determining the portion reliably representing a feature of said

speech waveform based on the first portion extracted by said extracting means

the second portion estimated by said estimating means (page 10.9.3, left column,

3rd full paragraph-right column, line 18) (e.g. The cited portion discloses the

syllabic nuclei boundary estimate and then extraction of stable regions of the

syllabic nuclei.)

It would have been obvious to one of ordinary skilled in the art at the time

the invention was made to have modified the determination of a reliable portion

of a speech waveform as taught by Lea in view of Mermelstein with the inclusion

of extracting stable regions as taught by Schmidbauer. The motivation to have

combined the references involves the ability to do further processing including

context specification and stress pattern of utterances (see page 10.9.1, left

column, 3rd full paragraph).

As to claims 2, 9, and 15, Lea in view of Mermelstein in view of Schmidbauer teach all of the limitations as in claim 1 above.

Furthermore, Lea teaches wherein said extracting means includes voiced/unvoiced determining means for determining, based on said data, whether each segment of said speech waveform is a voiced segment or not (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. Voiced and unvoiced determination is made.) of said waveform of energy distribution of the prescribed frequency range of said speech waveform on the time axis (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. In the cited section a prescribed frequency range is used and dips of energy define minimums.),

Furthermore, Mermelstein teaches the means for separating said speech waveform into syllables at a local minimum (see page 881, right column, 1st full paragraph, and Figure 1) (e.g. The minimum of Figure 1 is used to determine and segment syllable.); and

Furthermore, Lea teaches the means for extracting that range of said speech waveform which includes, in each syllable, an energy peak in that syllable within the segment determined to be a voiced segment by said voiced/unvoiced determining means and in which the energy of the prescribed frequency range is not lower than a prescribed threshold value (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is used to determine voiced and unvoiced segments. A frequency range for sonorant

energy is defined and since dips are located it is seen intuitively that maximums

will occur.)

As to claims 4, 11, and 17 Lea in view of Mermelstein in view of Schmidbauer

teach all of the limitations as in claim 1 above.

Furthermore, Lea wherein said determining means includes means for

determining, as a highly reliable portion of said speech waveform, a range

included in the range extracted by said extracting means, within the range of

which change in speech waveform is estimated by said estimating means to be

well controlled by said source (see Figure 2 and page 42.7.3, right column, 1st full

paragraph) (e.g. Form the figure, the syllables are detected and a range in time is

specified as seen ion the frames on the x-axis, "island of reliability") (e.g. It would

have been obvious to extract the frames corresponding to the extracted syllable

as defined by the timing in the Figure (e.g. Frames). Further, the use of a voice

detector as denoted in Lea will provide a range for voicing compared to unvoiced

segments.)

9. Claims 5, 6, 12, 18, and 19 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Lea in view of Mermelstein.

As to claims 5, 12 and 18, Lea teaches a quasi-syllabic nuclei extracting

apparatus for separating a speech signal into quasi-syllables and extracting a nuclear

portion of each quasi-syllable, comprising:

voiced/unvoiced determining means (see Figure 1, voicing decision) for determining whether each segment of the speech signal is voiced or not (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1, voicing decision) (e.g. Voiced and unvoiced determination is made.);

means for separating said speech signal into quasi-syllables (see Figure 1 syllabic nucleus detection) at a local minimum of time-distribution waveform of an energy of a prescribed frequency range of said speech signal (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. In the cited section a prescribed frequency range is used and dips of energy define minimums (The term quasi-syllable was interpreted to mean relating to a syllable.); and

means for extracting that range of said speech signal which includes energy peak in each quasi-syllable (see Figure 1, energy calculation and syllabic energy detector), determined by said voiced/unvoiced determining means to be a voiced segment and of which energy of the prescribed frequency range is not lower than a prescribed threshold value, as the nuclei of quasi-syllable (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is used to determine voiced and unvoiced segments. A frequency range for sonorant energy is defined and since dips are located it is seen intuitively that maximums will occur. Both the syllabic nucleus detection and voicing decision are interconnected.).

However, Lea does not specifically teach the minimum of a time distribution waveform.

Mermelstein does teach the use of a time distribution waveform for detecting local minimums (see Figure 1, and page 881, left column, sect. I, entire section) (e.g. The cited section uses a convex-hull to determine local minimum on loudness versus time waveform.)

It would have been obvious to one of ordinary skilled in the art at the time the invention as made to have modified the separation of speech signal into quasi-syllables as taught by Lea with the use of a time-distribution waveform as taught by Mermelstein. The motivation to have combined the references involves the ability segment of speech into syllable units (see Abstract) more effectively.

As to claims 6 and 19, Lea in view of Mermelstein teach all of the limitations as in claims 5 and 18, above.

Furthermore, Lea teaches wherein said extracting means includes means for extracting that range of said speech signal which includes an energy peak in each pseudo-syllable within the segment determined to be a voiced segment by said voiced/unvoiced determining means and in which the energy of said prescribed frequency range is not lower than a prescribed threshold value as the nuclei of quasi-syllable (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is sued to determine voiced and unvoiced segments. A frequency range for sonorant energy is defined and since dips are located it is seen intuitively that maximums will occur.). Furthermore, Mermelstein teaches the use of determining the peak of the loudness function in order to

determine the syllable boundary (see page 881, right column, 1st full paragraph
and Figure 1).

Allowable Subject Matter

10.   Claims 3 and 16 would be allowable if rewritten to overcome the rejection(s)
under 35 U.S.C. 112, 2nd paragraph, set forth in this Office action and to include all of
the limitations of the base claim and any intervening claims.

11.   Claims 7 and 13 would be allowable if rewritten or amended to overcome the
rejection(s) under 35 U.S.C. 112, 2nd paragraph, set forth in this Office action.

12.   The following is a statement of reasons for the indication of allowable subject
matter: None of the prior arts or combination there of teach the limitations as recited in
claims 3, 7, and 16 as that of "based on an output from said linear predicting means,
distribution on the time axis of local variance of spectral change" and "…means for
estimating, based on both … first calculating means and … second calculating means".
Most of the prior arts disclose the inclusion of the first calculating means.

Conclusion

13.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time
policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE
MONTHS from the mailing date of this action. In the event a first reply is filed within
TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to PARAS SHAH whose telephone number is (571)270-

1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m.

EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number

for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Paras Shah/
Examiner, Art Unit 2626

05/22/2008

/Patrick N. Edouard/
Supervisory Patent Examiner, Art Unit 2626